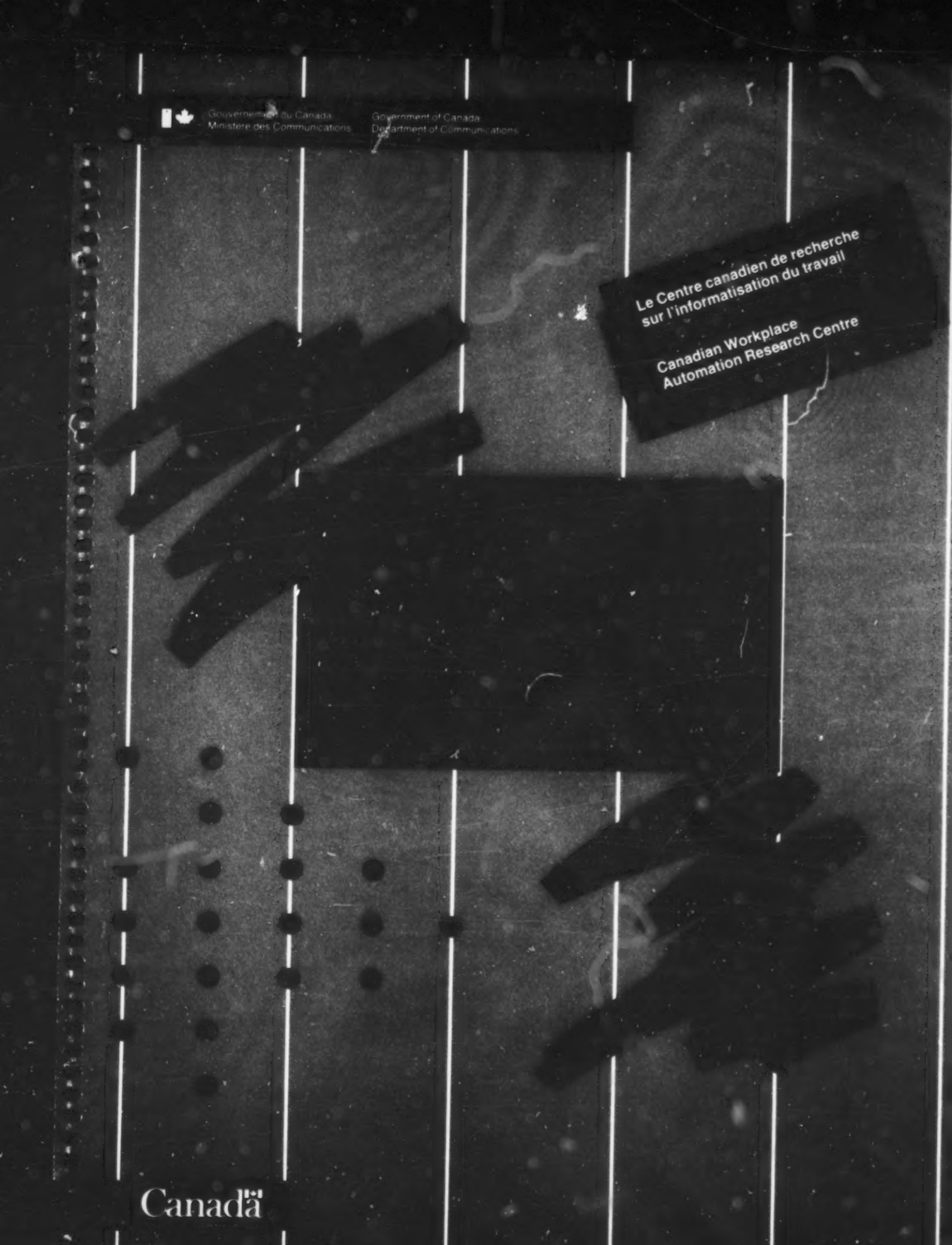
Gouvernement du Canada
Ministère des Communications
Government of Canada
Department of Communications
Le Centre canadien de recherche sur l'informatisation du travail
Canadian Workplace Automation Research Centre
Canada

THE TRANSLATOR'S WORKSTATION PROJECT

by

Elliott Macklovitch
Pierre Isabelle

Researchers at the
Canadian Workplace Automation
Research Centre

Communications Canada

Laval
October 1988

90-05453/

Copyright Minister of Supply and Services Canada 1988.

Cat no. Co28-1/22-1988E

ISBN 0-662-16497-0

The views expressed in this report are those of the authors only.

* Rapport complet en français aussi disponible.

PREFACE

The four papers in this collection trace the evolution of the translator's workstation project at the Canadian Workplace Automation Research Centre (CWARC). The first - "A Critique of the 1986 Socioscope Report" - was written shortly after responsibility for the project was transferred to the Advanced Technology Directorate at the CWARC, and highlights what we considered to be the weaknesses in the functional specifications that Socioscope had proposed in 1986. The second paper - "Update on the Translator's Workstation Project" - was distributed to all the translators and managers in the Translation Bureau of the Department of the Secretary of State; it consists of a brief resume of the project and a questionnaire that was intended to determine their priorities for a workstation. The third paper - "From Consultation to Specifications" - describes the results of our tabulation of the questionnaires and the implications of those results for the design of a workstation for Secretary of State translators. The actual functional specifications for such an installation of workstations are provided in the final paper of this volume. They in turn formed part of an RFP for technical specitications, the contract for which was eventually awarded to Odyssey Research Associates Inc. Odyssey's report containing the technical specifications for a translator's workstation is available from the CWARC as a separate publication.

A CRITIQUE OF THE 1986 SOCIOSCOPE REPORT

Elliott Macklovitch & Pierre Isabelle

Canadian Workplace Automation Research Centre

October 1986

I. Background

The idea of developing a specialized translator's workstation has been around for some years now. Within the Translation Bureau of the federal Department of the Secretary of State, the idea was first broached in studies conducted in 1979 on the advantages of providing translators with dedicated word processors, and on the possibility of augmenting these machines with on-line dictionaries and direct access to the Bureau's Terminology Bank.

Martin Kay of Xerox Corporation was probably the first to attract widespread attention to the notion of a translator's workstation. In an article written in 1980, entitled "The Proper Place of Man and Machines in Translation," he takes the position that given the current state of our knowledge about natural language, fully automatic machine translation is more or less a dead end, and that a far more cosntructive approach would be to provide the human translator with a set of automated aids. The idea of a translator's workstation is also discussed in the so-called Cognos Report of 1984-85, where it is suggested as a short-term development project within a new, national R&D effort in the area of automated translation.

From the start, the basic idea has always been to ergonomically integrate and adapt various forms of existing technology to the specific needs of the human translator, so as to allow him to perform routine administrative, research and text handling tasks more efficiently. The Translation Bureau's interest in the project derives essentially from this prospect of increased productivity: the workstation, it is hoped, will enable the Bureau to more effectively fulfill its mandate of providing translation services to all federal departments and agencies.

Communications Canada is interested in the project because the workstation represents an application of workplace automation, but also because of its commercial potential. The workstation is seen as a product which could find an interesting niche market both domestically and internationally, and which could therefore serve as a stimulus to the Canadian hi-tech sector. When Communications Canada created the CWARC in 1984, the project of developing a translator's workstation was included as part of its translation automation program.

On July 29, 1985, the CWARC and the Translation Bureau jointly issued an RFP, calling for a study that would culminate in a set of functional specifications for a translator's workstation. The contract for that study was awarded to the Ottawa-based firm of Socioscope Inc., which had enlisted a number of academics and specialists from the private sector to assist in the project. At the time the contract was awarded, there was a clear expectation in the two government departments that after this report had been submitted, thoroughly discussed and digested, it would be possible to

draw up a second set of technical specifications, on the basis of which a prototype could then be constructed and tested in the field. Indeed, the original RFP called on bidders to submit suggestions for an eventual pilot project involving the proposed workstation.

Socioscope submitted its report, entitled "Functional Specifications for a Translator's Workstation," on March 30, 1986. The report was read by a number of officials in both departments, and it seems fair to say that the response was not overly enthusiastic. The report was felt to be somewhat disconnected. No explicit link was made between a detailed task analysis describing a generalized "translation process" given in Section A and a long list of suggestions on various ways in which computerized aids could assist translators given in Section B. Moreover, the procurement timetable and technology assessment presented in Section C created the impression of a very ambitious project involving a substantial development effort. According to that timetable, for example, delivery of the suggested workstation would not begin before late summer of 1990. Socioscope was provided with feedback on the report and asked to correct at least the typos and errors of presentation that appeared in the original version. To ask for more than these superficial corrections, however, would have entailed extensive revisions to the entire report.

Since Socioscope submitted the final version of its report, scientific responsibility for the workstation project has been transferred from the Integrated Systems Branch at the CWARC to the Advanced Technology Branch. The purpose of this paper is to allow the authors, who work in the Automated Translation group of Advanced Technology, to describe the current situation with regard to the workstation project. More precisely, we wish to present the current state of our reflexion on the project, since, to the outside observer, it may appear that very little has been done over the last several months.

II. Current status of the workstation project

One action that we have undertaken over the last few months is to circulate the Socioscope report among a fair number of interested readers -- translators for the most part, but also to several people involved in private sector R&D. At the Translation Bureau, half a dozen section chiefs were asked to read the report and to put their reactions in writing, in the hope that these would provide us with some indication as to whether the Bureau's translation managers endorse the ambitious proposals advanced by Socioscope. As the reader may judge from their comments, which have been included in the Appendix to this paper, this does not appear to be the case.

There is at least one point on which most of those who read the Socioscope report do seem to agree, and that is that it would be extremely difficult, on the basis of this report, to proceed directly to draw up technical specifications for a prototype version of the translator's workstation. The basic reason has to do with the sorts of descriptions contained in those sections of the report that deal with functional

requirements (Section B.3) and functional specifications (Section C). Rather than providing a detailed description of just those automated functions that the workstation should be designed to perform, what these sections contain is essentially a long enumeration of suggestions on how many of the discrete tasks routinely performed by translators could be automated. No attempt is made to justify the advisability of automating this or that task, either in terms of potential improvements to the quality of the translator's working conditions or in terms of proven productivity gains that have been obtained through similar experiments in automation elsewhere. In short, there is no attempt to prioritize the numerous suggestions contained in the report. What we have here are not really functional specifications at all, but a hodge-podge of more or less useful suggestions, with the authors simply avoiding the really difficult questions of choice.

A related failing of the Socioscope report is that it does not examine currently available tools which could be directly integrated into a translator's workstation. This, in spite of the fact that Socioscope's original RFP proposal promised to provide an inventory of all relevant, available software. What's more, the final report argues on several occasions for the use of off-the-shelf software wherever possible. It is all fine and well to suggest that the workstation should incorporate on-line dictionaries, to take but one example. But are such dictionaries currently available, and if so, at what cost? More generally, the report does not indicate which components of an eventual workstation would have to be developed from scratch (and how long it would take to do so), and which components could reasonably be obtained off the shelf.

The issue of costs alluded to above is another major question that must be resolved before the technical specifications for a translator's workstation can be undertaken. Attempting to implement even half the suggestions contained in the Socioscope report would entail considerable development costs: this seems quite clear from the Technology Assessment provided in Section C. Presumably, these costs will eventually have to be recovered through the sale of the workstation. Before embarking on such a development effort, therefore, it would be highly advisable to examine the market potential for this type of product. To begin with, just how many professional translators are there in Canada? And how many of these would be prepared to spend (or could convince their employers to spend) x thousand dollars on a workstation -- whose productivity benefits have yet to be demonstrated?

Until at least some of these questions can be answered, we feel that it would be premature and perhaps even foolhardy to attempt to draw up the technical specifications for a translator's workstation.

III. Where do we go from here?

It may well be that the difficulty and complexity of the workstation project were originally underestimated, perhaps because it does not involve a substantial research component. (In fairness to Socioscope, the work statement of their contract may have been overly demanding as well, given the time and money allotted them). Be that as it may, the urgent question is what to do now in order to get the project moving again. One positive development is that the Technology Assessment division of Communications Canada has announced its intention to call for a market and industrial assessment study of computer-assisted translation systems. While we do not wholly endorse all the stated objectives of this study, it should nevertheless provide some basic information about Canada's translator population and its characteristics as a target market for a product such as the workstation.

All work needn't be suspended, moreover, until the results of this study are in. There would seem to be no reason why additional studies could not be concurrently undertaken in an effort to compensate for other of the above-mentioned weaknesses in the Socioscope report. In particular, we suggest that one such study be set up to gather information on existing tools (software and hardware) that might be integrated into a translato's workstation, including the cost of each potential component, its compatibility requirements and previous trials with it that could provide useful productivity data.

This study should take no more than a month for a consultant to complete. Using the information gathered, it might then be possible to organize a small-scale trial within the Translation Bureau, the goal of which would be to validate some of the very basic assumptions underlying the workstation project. For example: How many translators would be willing to abandon their dictaphones in favor of a keyboard and computer terminal that could offer them more than just word processing? Is it any more difficult for translators to adapt to a microcomputer running several application programs than it is for them to adapt to the sort of dedicated word processors that are currently being used in the Bureau? Do similar productivity gains occur? Are micros more flexible than dedicated word processors in allowing for direct telecommunication of texts between the translator and various clients with different word processing equipment? Do translators have a marked preference for any one of the numerous word processing packages currently on the market? How adequate for their needs are the spelling checker programs that several of these packages now offer? How useful do translators find commercially available software for building personal glossaries? Does a split-screen display on a standard micro terminal allow enough text in each window for the purposes of translation? Etc, etc.

In our view, a good deal of important information for the workstation project could be obtained from a relatively modest trial at the Translation Bureau. The basic arrangement would be for the CWARC to organize and conduct the study while the Bureau provided the translators and covered the cost of the material. It should be mentioned in passing that Bureau management is quite favorable to the idea of replacing their older word processors with microcomputers, and that several section chiefs have already expressed an interest in having their groups participate in an eventual trial of a prototype workstation.

Government of Canada / Gouvernement du Canada

MEMORANDUM **NOTE DE SERVICE**

TO / À

Monsieur Fernand Gobeil
Directeur
Technologie
DGPT

FROM / DE

Hélène Brisson
DSTM - Emploi et Immigration
DGOT

SECURITY - CLASSIFICATION - DE SÉCURITÉ	
OUR FILE – N / RÉFÉRENCE	
YOUR FILE – V / RÉFÉRENCE	
DATE	Le 12 août 1986

SUBJECT / OBJET

Rapport Socioscope - Poste de travail automatisé

Les personnes suivantes se sont réunis le 7 août 1986 pour discuter du rapport Socioscope sur un poste de travail automatisé : Ginette Côté, Maria Virjee, Pierre Danis, Henri Frickx et Richard Houde, ainsi que la soussignée. Nous avons discuté du processus de traduction qui y est exposé et du genre de poste de travail qu'on aimerait voir implanter.

Les observations sont présentées sous forme de points.

Commentaires généraux

1. Aucun projet du genre ne devrait être entrepris ni mené à terme sans que les questions de mobilier et d'environnement ne soient prises en compte. Le groupe considère la chose essentielle (voir en particulier les pages 95-98 avec lesquelles le groupe est d'accord).

2. Quoique le texte ait été révisé, il demeure pratiquement inacceptable du point de vue de la formulation et contient encore des contradictions et des erreurs.

3. L'échantillon des traducteurs interviewés (18 au total - privé et public) est trop faible pour être représentatif. Le BT devrait faire son propre sondage et interroger en particulier ceux qui utilisent des MTT ou des micro.

4. L'échéancier proposé est parfaitement irréaliste. C'est un horizon de planification beaucoup trop court.

5. Notre philosophie devrait être "Walk, Don't run"; nous devrions corriger nos lacunes avant de nous lancer dans l'aventure du poste de travail automatisé. L'ébauche de projet paraît grandiose pour un proche avenir tout en étant des plus intéressantes à long terme (au-delà de l'an 2000).

.../2

6. Nulle part, il n'est fait mention du soutien. Sans doute parce que l'hypothèse de travail sous-jacente mais non exprimée, c'est celle du traducteur autonome; or, il est faux de croire qu'un poste de travail automatisé éliminerait toute intervention du soutien. Ce n'est pas, en tout cas, ce que laisse entendre le rapport, qui ajoute la révision au nombre des interventions. Donc, il faut, au contraire, l'inclure dans toute étude et toute planification. Surtout que cette technologie transformerait certainement leur rôle. À étudier : la rentabilité de qui fait quoi.

7. La sécurité des documents et du poste mérite à peine un court paragraphe. C'est méconnaître l'importance de la question pour nous. À fouiller bien davantage.

8. Si nous sommes tous "branchés", qu'advient-il des entrepreneurs? Les répercussions d'une mécanisation généralisée du BT se feront sentir sur eux.

9. Nos clients seront sans doute très intéressés de suivre l'affaire: il faudra leur rappeler que le poste de travail automatisé existe pour répondre à nos besoins et non à leurs exigences.

10. Et enfin, les questions qu'il aurait d'abord fallu poser : qu'en sera-t-il du BT dans trois, quatre, cinq ans? Comment un tel projet s'insère-t-il dans la planification stratégique du BT? Et surtout des Opérations de traduction, usager éventuel?

Le processus

1. Le processus décrit existe bel et bien dans l'idéal, mais non dans les faits; les étapes ne sont pas suivies ou elles sont escamotées. Il est probable que les personnes qui travaillent déjà sur MTT organisent leur travail de façon beaucoup plus systématique. Une longue description qui ne révèle rien de nouveau.

2. Encore une fois, on ne distingue pas suffisamment la source des renseignements. La lecture du milieu laisse à désirer: elle manque de pondération. Qui plus est, on constate, et ce faisant, on répète des vérités connues depuis longtemps. Pas d'analyse sous un angle nouveau. À priori discutables (dictée, moyenne quotidienne de production, temps de recherche, par ex.). Les contingences propres au Multilingue (terminologie, par exemple) ne ressortent pas du tout.

3. Le BT devrait produire sa propre description du processus. À ce titre, l'analyse de la tâche faite par un sous-groupe du comité des usagers du SIO, il y a environ un an, serait des plus utiles. À examiner aussi, les études déjà menées au BT sur différents systèmes.

.../3

Le poste de travail automatisé

1. Le poste de travail automatisé est un outil parmi d'autres. Les gestionnaires devraient pouvoir se réserver le choix des outils. Le poste n'est pas la solution universelle. Pourquoi vouloir à tout prix écarter la dictée? Certains seront plus productifs s'ils continuent à le faire. Faut-il penser que tous les TR seraient dotés d'un poste? Ce ne serait pas nécessairement rentable. Qui plus est, si l'on maîtrise le fonctionnement de l'outil, on est dégagé pour la traduction proprement dite (qualité).

2. L'obligation de se servir d'un clavier pourrait être un élément dissuasif. Il semble en effet que dactylographier soit un obstacle plus grand pour le TR que le fonctionnement d'un système.

3. Le BT devrait demander aux universités de s'assurer que les futurs traducteurs maîtrisent une méthode de dactylographie et une méthode de dictée.

4. Le poste idéal devrait d'abord et avant tout donner accès à des sources d'information: TERMIUM, banque de textes. Le TR devrait aussi pouvoir établir son propre dictionnaire. Le "split-screen" serait très utile à l'étape de l'auto-révision. (La qualité et la taille de l'écran importent beaucoup dans ce cas.)

5. La possibilité de communication électronique avec les clients est intéressante, mais il ne faudrait pas oublier les exigences administratives; donc, pas de communication directe de texte du client au TR et vice-versa. L'hypothèse selon laquelle tout le monde ou tous les systèmes d'information sont automatisés est fausse: nous avons et nous aurons d'énormes problèmes de compatibilité. La possibilité de remettre un texte prêt à imprimer est cependant très intéressante du point de vue de la qualité du service.

6. Pourquoi ne pas envisager dès maintenant la possibilité de postes TR fonctionnant grâce à la reconnaissance de la voix?

7. On devrait organiser une démonstration des composantes décrites dans la seconde partie du rapport (écran plasma, etc.), et des composantes déjà sur le marché (dictionnaire électronique).

c.c.: G. Asselin
M. Robichaud
G. Côté
M. Virjee
P. Danis
R. Houde
H. Fricks

UPDATE ON THE TRANSLATOR'S WORKSTATION PROJECT

Elliott Macklovitch
March, 31 1987

1. What is the translator's workstation?

The translator's workstation, as we use the term, currently exists only as a project. The aim of this project is to provide the human translator with a set of tools or computerized aids which will improve his workig conditions and render him more productive. Basically, the idea is to automate certain simple, repetitive tasks, thereby allowing the translator to concentrate on the more difficult jobs that genuinely require his specialized skills.

The following are some of the tasks which a workstation could help a translator perform more efficiently: consulting terminology records; detecting spelling errors; comparing different versions of the same text; counting the words in a text; and of course typing, editing and formatting texts.

All these functions will likely be integrated and implemented on a microcomputer. The workstation may thus be viewed as the successor of the dedicated word processors currently being used by some translators in the Bureau. Unlike these word processors, however, a microcomputer can run a wide variety of application programs. The workstation to be delivered by this project will offer much more than word processing and should be useful even for those translators who would rather not type their own translations.

2. Who is responsible for the workstation project?

The Canadian Workplace Automation Research Centre (or CWARC) is responsible for coordinating work on this project.

The CWARC is a new Communications Canada laboratory, located in the Montreal suburb of Laval. Its mandate is to promote research and development on computerized technologies that are intended to increase productivity and improve working conditions for a whole range of jobs currently performed on the labor market. Translation is one of the areas that has been targeted as a priority by the CWARC, and the Centre has recently established an R&D team in machine-aided translation, under the direction of Pierre Isabelle. The Bureau has seconded Elliott Macklovitch, its machine translation project officer, to join this team at the CWARC.

The development of a specialized translator's workstation is one of this team's major projects, one in which the Bureau clearly has a special interest. Within the Bureau, Anne Bordé, acting director of Technology, is responsible for ensuring liaison with the CWARC.

3. What is the current status of the workstation project?

In July 1985, the CWARC and the Translation Bureau jointly issued an RFP calling for a study on the functional specifications of a translator's workstation. That contract was awarded to the Ottawa-based firm of Socioscope Inc., which submitted its final report on March 30, 1986. Although Socioscope did consult a certain number of translators during its study, the officials in both departments who read the final report were hesitant to endorse Socioscope's proposals for a translator's workstation. Socioscope's vision of the workstation is an extremely ambitious one, and involves integrating a wide range of leading-edge technologies, many of which will not become affordable before the of the decade. Moreover, the authors of the report fail to establish any priorities among all the functions that they propose to automate, and do not examine any existing products which could well form part of a first version of the workstation.

The CWARC's workstation project is more modest than the one advanced by Socioscope and, in our view, far more realistic. Our goal is to produce an initial prototype within one year. In order to meet this deadline, we must first determine which functions it would be most cost effective to automate, and among these, which can be automated using off-the-shelf software and hardware. To help answer questions like these, the CWARC has recently undertaken a small survey of freelance translators and translation departments in the private sector, in an effort to learn what equipment translators are currently using in the most automated services and hopefully benefit from their experience.

4. Will the Bureau's translators be consulted on the design of this new workstation?

Most certainly! As mentioned above, we are presently in the process of preparing the functional specifications for a first version of the workstation. The crucial question here is the following: among all the functions and tasks involved in translating or providing translation services, which should be automated first? Those responsible for the project need to know the views of the Bureau's translators and managers on this question.

To help us obtain those views, we have prepared a short questionnaire which we have attached to this document and are asking you to complete. It is important that as many of the Bureau's translators and mangers as possible respond to this questionnaire, so that the people at the CWARC can attain a clear understanding of your needs and opinions.

This questionnaire certainly does not exhaust all possibilities for a translator's workstation. Should you have other suggestions or comments to make on the project, please do not hesitate to add them to the bottom of the questionnaire.

5. What are the next stages in the workstation project?

If all goes as planned, the functional specifications for the translator's workstation should be drafted by the end of April 1987. Those responsible for the project will then prepare the technical specifications which will allow for the production of the first prototype.

This stage will probably involve a certain amount of programming and development work, since it is unlikely that we will find commercial software corresponding to all the desired functions for the workstation. It will also be necessary to take into account other projects currently underway at the Bureau that are certain to have an impact on the workstation, such as the transfer of the terminology bank onto optical disk. Nonetheless, we are confident that we will be able to produce an initial version of a translator's workstation before the end of 1987-1988. A number of samples of this prototype will then be placed in the hands of Bureau translators so that they can test it in an operational environment and make suggestions for its subsequent improvement.

TRANSLATOR'S WORKSTATION PROJECT
QUESTIONNAIRE

Name(optional): _____ Section: _____

Are you a translator ____ a reviser ____ or a manager ____?

Do you currently use a word processor or a microcomputer to carry out your work at the Bureau?

yes ____ no ____ occasionally ____

This questionnaire lists computerized components that may eventually form part of a translator's workstation. Kindly indicate your opinion on the importance for the workstation of each of the listed functions by assigning it a rating of 1 to 3:

1. to indicated that in your view this function would only be minimally useful as part of the workstation;
2. to indicate a function that could at times prove useful, without being indispensable;
3. to indicate a function that should be included as a priority in the first version of the workstation.

A. A word processing program:

1. that can display all the accented French characters, as well as other symbols (e.g. technical and mathematical symbols, Greek characters, even other alphabets); _____
2. that allows for the easy editing of columns and tables; _____
3. that incorporates a split-screen display for source and target text; _____
4. other important functions (aside from those that are standard in most word processors): _____

_____

_____

B. Writing aids, such as:

1. a spelling checker; _____
2. a segmentation program that splits words only at syllable boundaries at the end of a line. _____

C. Tools for terminological research, such as:

1. a sort utility that can produce an alphabetical list of all the words in a text (with their frequency) and display each form in context; _____

2. a program for creating, editing and accessing personal terminology records; _____

3. a section-wide terminology file, accessible from each workstation; _____

4. direct access to the Bureau's terminology bank; _____

5. a dictionary of synonyms (or thesaurus), resident in memory and directly accessible within the word processing program; _____

6. standard dictionaries, resident in memory. _____

D. A modem and telecommunications program:

1. to receive source texts and deliver their translations electronically; _____

2. to query remote data banks. _____

E. Other utilities, such as:

1. a program that could compare two versions of a text and indicate the differences or updates; _____

2. a word counting program; _____

3. a personal agenda (or work planner) to enter the date a text was received, the date it is to be delivered, the time spent on it, etc; _____

4. a format conversion programsm (to facilitate the exchange of text files prepared under different word processing programs). _____

F. Other suggestions for the workstation, or comments: (use the next page if necessary)

_____

_____

_____

Once this questionnaire is completed, please send to:

Anne Bordé
Technology Directorate
Translation Bureau
Terrasses de la Chaudière, #7F13
Hull (Quebec)

THE TRANSLATOR'S WORKSTATION PROJECT:

FROM CONSULTATION TO SPECIFICATIONS

Elliott Macklovitch

June 11, 1987

In April 1987, a short document describing the background and current status of the translator's workstation project was distributed to all the translators in the federal Translation Bureau, along with a two-page questionnaire. The questionnaire enumerated a list of components which might be incorporated into a workstation, and asked Bureau employees to indicate their opinion on the importance of each component. The document stressed that this list of components was not to be seen as exhaustive and invited translators to append their comments and suggestions to the questionnaire. In all, more than 400 people replied to the questionnaire, and many of these did include comments and suggestions. This represents a response rate of over 50%, which under the circumstances is extremely good. (Owing to delays at the printers, some translators did not receive the questionnaire until after the date they were asked to return it.)

In this report, I present the results of my tabulation of the participants' responses and then discuss what we may infer from this large-scale consultation for the design of a translator's workstation.

Questionnaire results

In the enumeration that follows, the order of presentation reflects the respondents' priorities regarding the functions proposed on the questionnaire.

1. Direct access to the Bureau's terminology bank

Accross the Bureau, this was the function judged to be by far and away the most important. Part of the reason can probably be traced to the problems of access experienced by translators in the months preceding the survey. Another reason is that for those translators who do not intend to abandon their dictaphone for a keyboard, the main interest of the workstation lies in its terminological facilities.

Only the multilingual translators and those in the technical and scientific sections failed to accord this function the highest score, and this for reasons that seem relatively transparent.

2. A section-wide terminology file

A large number of translators, are of the opinion that it is extremely important for the workstation to include such a facility. At least one section in the Bureau has developed this sort of interactive terminology file. Some respondents noted that there might need to be several glossaries, one per subject area.

Only the multilingual and English translators rated this function lower than second.

3. A personal terminology file

The main advantage of a section-wide terminology file is that it allows each member of the group quick access to the results of his colleagues' research. A personal terminology file, on the other hand, simply automates a manual process, which may explain why this function scored significantly lower. In all probability, however, the personal and section-wide glossary programs will form part of a single terminology management facility.

4. A spelling checker

There is little doubt that a high performance spelling checker could ease the burden of final proofreading, especially for minor typos that often pass unseen. Very efficient checkers are available for English, but their French counterparts still need some development.

5. **A word processing program that allows for the easy editing of columns and tables.**

Beyond a number of basic functions that must be available, the complexity of the workstation's word processing software depends in large part on the extent to which translators are responsible for delivering a final version of their target text. It would probably be more cost-effective if the translator's responsibility were limited to producing a simple draft, with complicated formatting operations being handled by the support staff. This is also the arrangement that translators would seem to prefer. Nevertheless, the high score attributed to this question indicates that translators would appreciate a program that would allow them to handle columns and tables with a minimal of gymnastics.

6. **A file comparison program (to detect updates)**

Although the client is theoretically supposed to indicate updates, often he does not do so. Here is a perfect example of the sort of time-consuming task for which the computer is ideally suited - provided, of course, that both versions of the text are available in machine readable form.

7. **A segmentation program**

Here again, the importance of this function depends on whether the translator is responsible for delivering a "pretty" text.

8. **A word processing program that can display technical and mathematical symbols, Greek characters, foreign alphabets (in addition to all French characters in upper and lower case)**

   Not surprisingly, the needs of different translation services vary. Translators in scientific and technical sections ranked this component fourth in importance, just after direct access to Termium, while the multilingual translators ranked it first.

9. **A word processing program that can display source and target text side by side**

   In retrospect, this function probably should have been described differently. For translators who are liable to spend many hours a day staring at a monitor, its resolution, the physical size of the characters and the capacity to display a full page of text are probably more important. As one respondent commented, translators frequently need to browse through their text, and dividing an already limited display in two could be more of a hindrance than a help.

10. **A modem and telecommunications program to query remote data banks**

    Here too one must distinguish between the usefulness of various data bases for translators and the technical means of accessing those banks. Several translators who were aware of the CD-ROM project suggested consulting other data bases by means of the same technology.

11. Electronic versions of various types of dictionaries

If the reaction of the respondents was rather lukewarm, it may be because they see little advantage in gaining two or three seconds here and there by automating reference books that are already sitting on their desks.

12. A modem and telecommunications program to receive and deliver texts electronically.

Currently, it is the support staff that is responsible for the receipt and delivery of texts. Comments on this point were quite explicit: translators do not want this to change with the introduction of the workstation.

Same comment for the word count.

13. An automated "Bescherelle" for French translators

Same comment as for function 11 above. Qualified translators rarely need to consult a Bescherelle. However, the question only applied to French translators, which accounts in part for why it scored so low.

14. A format conversion program for text files prepared under different word processing programs

The growing demand for translations delivered in machine readable form is creating serious compatibility problems for the Bureau, owing to the diversity of equipment in various government departments. Until industry-wide communication protocols are implemented, one alternative

to the rather costly and time-consuming solution provided by Keyword would be to incorporate file conversion tables directly into the workstation. Once again, however, translators feel quite strongly that this activity should be left to support staff.

15. **A sort utility that could alphabetically list all the words in a text (with their frequency), and recall any given form in context**

The WANG OIS system currently used in the English and technical sections offers a similar utility, which translators apparently find quite useful. Perhaps part of the reason for this question's low rating was that the manner in which it would operate was not made sufficiently clear.

16. **A personal agenda for text tracking**

Seeing that this type of information is already entered in the Operational Information System (OIS), several managers suggested that the workstations be interfaced with OIS.

Functional specifications

From the priorities expressed on the questionnaire, it seems quite obvious that the workstation cannot be implemented as a standalone installation; the sharing of information, software and peripherals clearly calls for a network of linked stations.

It is also apparent that the needs of different groups of translators - especially for word processing - are not homogeneous across the Bureau (as discussed in sections 5 and 8 above). Hence the workstation must be modular in design, and assembled "à la carte".

Pratically speaking, here is how a unit equipped with workstations might operate. Bear in mind that what follows is an illustrative example and that other types of work arrangements (spliting a text between several translators or units, revision or autonomy, etc.) could be accomodated as well.

1. Receipt of the source text

Ideally, the source text is received in machine readable form, either on diskette or via telecommunications. From her workstation, the section clerk submits the text file to a word count program, enters administrative information in the OIS, and then electronically forwards the text to the assigned translator (perhaps with an accompanying hardcopy).

2. Searching through previous translations

On the questionnaire, several translators suggested that the network should provide an index of all texts previously translated in the unit. This index could be searched under different fields (eg title of the document, client, delivery date) and would refer the translator to the file cabinet where a copy of the previous text and its translation is stored. A step further would be to archive all texts in machine readable form, so that when the translator located a pertinent text in the index, he could submit it and his current text to a file comparison program.

3. Terminological research

From his own workstation, the translator queries Termium (either on line or on the CD-ROM) after having produced, if he so desires, an alphabetical list of all the forms that occur in the text, along with their frequency. At any time during the translation process, he can also consult or update his personal glossary or the section's without having to exit from his word processing program.

4. Drafting the target text

Certain translators suggested that the workstation should include the means of automatically converting recorded dictation into a word processing file. Unfortunately, automatic speech recognition has not yet advanced to the point where this is a practicable option. Until then, the best the workstation can offer is a choice of word processing software that satisfies the characteristics mentioned above and is capable of displaying a full page of text on a high resolution monitor.

5. Revision, proofreading and delivery

If the translator is not autonomous, he forwards his translation to the revisor, electronically or in printed form. All workstations in the network will have access to two types of printers: one for rapid, draft quality and another for final, letter quality. Translators should be able to control the sequence of texts to be printed by means of a print queue.

Once the text has been revised, corrections are entered either by the translator or by the support staff, who will also take care of the final formatting. The text is then proofread one last time with the help of the spelling checker, to detect any remaining typos.

The clerk enters relevant data in the OIS; photocopies source and target texts, or archives them electronically after having updated the index of past translations; and then delivers the translation to the client, either on paper or via telecommunications.

Summary of the functional specifications

The functions proposed for the first prototype of the workstation are as follow:

1. Hardware

   a) a network of linked workstations;
   b) each equipped with a high resolution monitor that allows for a full-page display;
   c) and sharing draft and letter quality printing facilities;
   d) a telecommunications port on the support staff's station;
   e) access to Termium on line or on CD-ROM.

2. Software

   a) a choice of word processing packages;
   b) a spelling checker;
   c) a terminology management program;
   d) a sort and recall program;
   e) a file comparison program;
   f) an index of previous translations;
   g) a word count program (for the support staff's station).

FUNCTIONAL SPECIFICATIONS
FOR AN INSTALLATION OF TRANSLATORS' WORKSTATIONS
IN THE TRANSLATION BUREAU OF THE DEPARTMENT OF THE SECRETARY OF STATE

A. Hardware

1. A network of linked workstations

a) On a questionnaire that they received in April 1987, translators in the federal Department of the Secretary of State clearly articulated the need to share certain automated tools and resources, e.g., an interactive section-wide terminology file, which cannot be easily accommodated by a standalone workstation. The installation hereafter described is intended to serve working groups of up to 30 persons. It must provide for the linking of individual translator's workstations and for the easy exchange of electronic information among these workstations. (At times, we refer to this installation as a network, though in no way is this intended to preclude the possibility that it might be implemented on a minicomputer as well as a LAN: cf. #1(g) below).

b) Other important advantages offered by a network include the sharing of expensive peripherals (e.g. a laser printer), the exchange of files among translators and revisers, and electronic messaging.

c) One designated workstation in the network shall be equipped with certain utilities, e.g., a telecommunications port, and will normally be operated only by support staff. This workstation may also require a large capacity storage facility for the electronic archiving of reference texts -- or, at minimum, a special indexing facility for completed translations, making possible sophisticated queries from any workstation in the network in order to help locate useful reference texts. (See section B.4 below.)

d) The workstations must be modular in design, permitting the replacement or addition of a given component while retaining the advantages of a common overall design across the Bureau. This

requirement is most apparent with regard to the workstation's word processing component. More and more Secretary of State clients are requesting that their translations be delivered on diskettes in a particular word processing format, which they can use directly without having to worry about format conversions. The problem is that different clients employ different word processing packages. If the Bureau is to satisfy this growing demand without massive recourse to expedients like Keyword, then the workstation will have to be able to run a number of the most popular word processing packages.

This need for modularity is reinforced by the fact that new components will have to be added to the workstation as they become available or as their applicability to the translation process is identified. On the above-mentioned questionnaire, for example, translators expressed the need for a file comparison program that could rapidly identify updates or modifications in two electronically supported versions of a text. To our knowledge, no such program currently exists which is specifically designed for word processing files. It is crucial that the architecture of the first version of the workstation be open enough to allow for the easy integration of such a program when it does become available. (The same applies to an efficient grammar checker or other writing aids that could usefully be integrated within the workstation.) Furthermore, the workstation must be endowed with sufficient memory to allow it to run additional application programs.

e) There are a number of other characteristics which it would be desirable for the workstation network to incorporate. The network should offer a degree of integration and coherence among its various components. At a minimum, this should consist of a menu or some other simple means of selecting among the available programs. More importantly, all the components must be able to function harmoniously together: the commands of one program, for example, should never

block or interfere with those of another. The need to share common resources should not prevent individual translators from maintaining personal files and controlling other users' access to those files. Translators should, whenever possible, be relieved of the chore of having to make precautionary backup copies of their files. It would also be preferable to be able to dialogue with the network's operating system in either English or French; and of course the same applies to all the system's other programs. And finally, the users of the network must never be completely immobilized or prevented from working because of a breakdown on a single station or network server.

f) There would appear to be at least two different ways in which such a network of workstations could be implemented: either by providing each translator with a terminal linked to a minicomputer; or by providing each translator with a microcomputer that is linked to other microcomputers in a local area network. The selected consultant will be required to justify the choice of one or other of these options, or propose alternate solutions that compare favorably with these, in terms of cost, desired modularity, coherence of the system's various components, failsafe characteristics, and other advantages.

2. A high-resolution monitor capable of displaying a full page of text

a) For translators who are liable to spend many hours a day staring at a monitor, the actual physical size of the screen, its resolution and the size of the characters it displays are very important ergonomic considerations. Furthermore, translators frequently need to refer to preceding and following paragraphs in the document they are drafting, and many find that standard 24-line monitors do not allow for the display of sufficient context. The workstation shall therefore provide for a high-resolution monitor, capable of displaying a full page of text without reducing that text to miniature dimensions.

b) On the other hand, the Translation Bureau is not frequently called upon to deliver camera-ready copy and most translators only have to handle text. A monitor for the workstation that offered sophisticated graphics or color capabilities will not therefore be required.

3. Draft and letter quality printing facilities

a) Each network of workstations shall include at least two printers (more, if the size of the section warrants it): one providing final, letter quality and a second providing rapid, draft quality as well as near-letter quality (in the event that the other printer breaks down). Both types of printers must be able to print the full range of characters that can be displayed on the workstation's monitor.

b) Translators should be able to control the sequence of texts to be printed on either printer from their workstation by means of a print queue.

c) The selected consultant will be required to justify a choice of printer types and models in terms of quality of the printed output, reliability, speed, flexibility, cost, etc.

4. A telecommunications port

a) Each network of workstations shall include a telecommunications port to allow for the electronic exchange of source and target documents with clients and other translation sections. This telecommunications port will normally be operated by support staff, who will continue to be responsible for the receipt and delivery of texts in the section.

b) Only two or three sections in the Translation Bureau currently employ a modem for this purpose: the vast majority of source texts continue to arrive on hard copy. On the other hand, more and more clients are requesting that their translations be delivered on diskettes. There are few requests for delivery via telecommunications, perhaps because this frequently occasions loss of French accents and certain formatting information.

c) There would appear to be at least two different ways in which such a telecommunications port could be implemented: either by installing a modem and a telecommunications program on a designated workstation, or by providing the network with electronic mail facilities. The selected consultant will be required to justify the selection of one or other of these options in terms of ease of use, reliability, cost, etc., and make specific proposals as to the implementation the selected option.

5. Direct access to the Bureau's Terminology Bank

a) The Terminology Directorate of the Translation Bureau is about to conduct an operational trial of its Terminology Bank (frequently referred to as Termium III) on CD ROM. The Toronto-based firm of Reteaco Inc. has been awarded a contract to customize its FIND-IT interrogation software to the needs of translators querying the Bank for terminological equivalents. For this trial, FIND-IT will be installed on an IBM PC/AT.

b) Translators would like to be able to access the Bureau's Terminology Bank directly from their workstations. Ultimately, they would hope to be able to access the Bank without having to exit from their word processing program. The selected consultant will be asked to specify precisely how this sort of integration between the workstation network and the Terminology Bank on CD ROM can be achieved.

B. Software

1. Word processing software

a) For translators who will spent most of their time at the workstation actually drafting texts, the word processing software will probably be the workstation's single most important component. There are a multitude of commercial word processing packages already on the market: the problem is to select which one or ones best meet translators' needs.

b) It is quite clear that the word processing needs of different groups of translators are not homogeneous across the Translation Bureau. Translators in the technical and scientific sections, for example, require special mathematical symbols for their texts more frequently than do their colleagues in other sections; they probably translate more tables and figures as well. Multilingual translators cannot consider a workstation that does not display Spanish accents, German consonants, or Cyrillic, Greek or Japanese alphabets. For most other Bureau translators, on the other hand, a word processing program only has to handle the English and French character sets, and perform a small number of basic operations in a manner that is simple and easy to learn.

c) Here again, there would appear to be two different ways of responding to the situation: either every workstation retains a number of alternate character sets in memory, allowing the user to redefine his keyboard in maximally simple manner; or the workstation provides particular groups of translators with specialized word processing software, tailored to their special needs, retaining for all other translators in the Bureau one or more "standard" word processing packages. The selected consultant will be required to justify the choice between one or the other of these options, or propose alternate solutions.

d) As mentioned above, a standard word processing package must be able to perform a relatively small number of basic text editing operations, such as: the insertion, deletion, movement and copying of various textual units; search and replace; rapid cursor movement and scrolling; the "undoing" of prior modifications; underlining; centering; etc. The software should also provide a well-structured tutorial and easily accessible help screens in English and French.

e) The selected word processing software must be able to take advantage of the full-page monitor, i.e., the two must be compatible. It would be preferable, moreover, if the translator had the option of displaying two word processing files on the screen simultaneously, and of easily copying meterial from one window into the other.

f) All word processing software must support the full range of French characters, including upper and lower case accents.

2. A user-updated, local terminology management program

a) Translators may consult the Bureau's Terminology Bank, but generally speaking, they cannot directly add to its contents without the intervention of a bona fide terminologist. For their own specialized or client-dependent terminology, many sections maintain a cardex, to which all translators in the section are encouraged to contribute. Though its standards are less exacting than those of the central bank, this cardex often proves extremely helpful and translators would definitely like to see it automated.

b) Each workstation shall therefore have access to a local terminology management program which allows the translator to create and maintain his own personal terminology records. This program must be fully compatible with the workstation's word processing software in at least two distinct ways. First, the program should allow the

translator to create and consult terminology records without having to exit from his word processing document. And second, in creating or amending terminology records, the translator should be able to enter accented French characters using the same keys as he does in his word processor. In addition, the program should allow the translator to cut a selected equivalent out of a retrieved record and insert it into his document with a minimum of keystrokes.

c) The terminology management program must also allow for the merging of each individual's personal records into a section-wide terminology file. This will be the file that translators consult when they want to benefit from the results of their colleagues' research, although they should not be able to modify it directly without going through the section's terminology coordinator. The program shall ensure that new records added to a personal or section file are automatically sorted in correct alphabetical order.

d) The format of the records proposed by the program shall be free enough to allow translators to insert -- in addition to language pair equivalents -- comments, sources, the date the record was created and any other information deemed useful.

3. A spelling checker

a) There is little doubt a high-performance spelling checker could be of considerable help to translators by easing the burden of final proofreading, especially in detecting the minor typos that often pass unseen. Each workstation shall therefore include a spelling checker for English or French, depending on the target language for which it is used.

b) Flexible spelling checker programs, which allow the user to add new entries to their dictionary and specify text-specific terms that the

program will ignore, are currently available for English. On the other hand, the French spelling checkers offered with the most popular word processing programs detect so many false errors as to be practically unusable. In part, the problem may be due to the fact that French is more highly inflected than English, requiring the entry of many more inflectional variants for each base form. However, new spelling checker programs are appearing on the market all the time and some of these may incorporate a French morphology component.

c) The selected consultant will be required to review the literature in order to locate evaluations of the best spelling checker programs and determine which (if any) satisfy minimal performance criteria to merit testing in the first version of the workstation.

4. A system for the archiving and/or indexing of previous translations or reference texts

a) Each section in the Bureau organizes and maintains its own file system of past translations. Before undertaking a new text, translators routinely consult these files to see if they can find a similar (or identical) text. This search through the files may take considerable time and does not always prove fruitful. Translators will often abandon the search, frustrated at not having located a pertinent text which they know that the section has previously translated.

b) If the classification hierarchy inherent in the section's filing system were simply transferred to an indexing facility on the workstation configuration, in a way that allowed information on completed translations to be easily added, this might considerably facilitate the location of relevant reference texts. Each translator could then consult this file from his own workstation. Initially, a

successful search need provide no more than a reference to the exact location of a copy of the text (on paper or diskette) in the section's files. Ultimately, however, one would like to be able to use such a facility in order to archive electronic copies of source and target texts on a large capacity storage medium.

c) Documentary or textual database management systems which allow for more sophisticated indexing and search techniques are no doubt available on the market. However, it must be remembered that each section in the Bureau will not have access to the services of a professional documentalist and that the translators and support staff will be responsible for managing such a system entirely on their own. The selected consultant will therefore be required to propose ways of implementing a section-wide index of completed translations in a manner that will facilitate the location of reference texts without being too complicated to be operated by the translators themselves.

5. A word counting program

a) The number of words translated is an extremely important production statistic in the Bureau, both for management and translators. All incoming texts must be counted, and last year they amounted to about 290 million words. To assist in this enormous task, the Bureau relies on specialized word counting machines, but these can only process high-quality printed copy that contains no graphics. The word counts for a great many texts must still be done manually by support staff.

b) As mentionned above, the workstation network will include a telecommunications port, operated by support staff and used for the electronic exchange of source and target documents with clients. For the source texts that arrive in the section in electronic form, this workstation shall also include a rapid and reliable word counting

program, so that the support staff is relieved of at least part of this tedious work. Translators do not want to be responsible for counting the number of words in incoming texts and, while some might still appreciate having such a facility on their workstation, this is not an essential requirement for the first installation.

6. An alphabetical sort program

a) Another utility which translators will find quite useful, especially for updating their personal or section-wide terminology files, is a sort program that produces an alphabetical listing of all the forms in a text along with an indication of their frequency. Such a utility is currently offered with a number of application programs.

b) As far as the workstation is concerned, it is important that this program be able to handle accented French characters and sort them in correct alphabetical order, such as one finds in a standard French dictionary. And of course, the program must be maximally simple to operate. Translators may find it difficult to operate a workstation composed of a number of disparate application programs if each demands its own elaborate repertoire of command sequences.

C. Future versions of the translator's workstation

1. We emphasize again that the functional specifications thus far described are for the installation of a first prototype of a translator's workstation. As was mentioned in section A.1.d, this workstation must be modular in design, to allow for the addition of new components as they become available. Where no satisfactory commercial product is available for a given component, the CWARC or the Secretary of State may undertake or sponsor appropriate development work. In this section, we briefly enumerate additional components which we would eventually like to see form part of future versions of the workstation. Some, like a file

comparison program, may be realized via short-term development projects, while others, like fully automatic MT, are the subject of long-term research and will probably not be available for many years.

2. A file comparison program (cf. A.1.(d) above)

3. An optical character reader

For the foreseeable future, most of the source texts which the Bureau receives will continue to arrive in paper form. For texts that contain many tables or elaborate formatting, it would be convenient to be able to import the formatting information directly into the target text by means of an optical character reader: the translator could then type his translation over the electronic copy of the source text. Another useful application for an OCR would be for reference texts that translators want to consult but which only exist in paper form. In order to take full advantage of a utility like the above-mentioned file comparison program, such reference texts would need to be converted into machine readable form.

4. Interface with the Operational Information System (OIS)

The Bureau has recently implemented an OIS system to help track translation requests and maintain statistics on translation activity in each of its many sections. Several section chiefs have indicated that it would be important to provide for an interface between the planned workstation installation and the AT compatibles on which the OIS currently runs.

5. Automatic voice recognition

The majority of translators in the Bureau dictate and many have no intention of giving up their dictaphones to enter their translations directly on a keyboard. For translators with limited typing skills, the main interest of the workstation which we have described lies in the improved terminological

and research facilities it will offer. Their interest in the project would no doubt increase if the workstation included the means of automatically converting their recorded dictation into an electronic word processing file, which they could then revise on the workstation.

6. Machine translation

When fully automatic machine translation systems become available which are capable of producing output that is of sufficiently high quality for translators to post-edit cost effectively, it would be natural to link such systems to the workstation network. A translator who received a text which he thought the system was capable of handling could then send it off to the machine translation module and later post-edit the waw output on his workstation. Other texts which did not correspond to the system's capabilities would be translated "humanly", using all the facilities the workstation puts at the translator's disposal.